24238

APPLICATION

FOR

U.S. PATENT

TITLE: **METHOD FOR ASSEMBLING PCR FRAGMENTS OF DNA**

INVENTORS: **GEORGE NELSON BENNETT, MARY LOU HARRISON**

CERTIFICATE OF MAILING BY EXPRESS MAIL 37 CFR §1.10

"EXPRESS MAIL" Mailing Label No.: ER771104776US

Date of Deposit: October 31, 2003

I hereby certify that this paper, including the documents referred to therein, or fee, is being deposited with the U.S. Postal Service "Express Mail Post Office to Addressee" service under 37 CFR §1.10 on the date indicated above and is addressed to: Mail Stop Patent Application, Commissioner for Patents, P. O. Box 1450, Alexandria, VA 22313-1450.

Type or Print Name: Bryan Lopez

Signature: Bryan Lopez

METHOD FOR ASSEMBLING PCR FRAGMENTS OF DNA

PRIOR RELATED APPLICATIONS

[1] This patent application claims priority to U.S. Provisional Application Serial No. 60/422,807, filed on October 31, 2002, the disclosure of which is incorporated by reference in its entirety herein.

FEDERALLY SPONSORED RESEARCH STATEMENT

[2] Not applicable.

REFERENCE TO MICROFICHE APPENDIX

[3] Not applicable.

FIELD OF THE INVENTION

[4] A process for assembling a series of DNA fragments generated by PCR into an ordered circular arrangement for replication and genetic work in cells. The method uses the 3' overhangs created by excision repair enzymes to direct the sequential ligation of PCR fragments on a solid scaffold. The resulting assembled DNA is removed with a site specific recombinase.

BACKGROUND OF THE INVENTION

[5] Current methods of manipulating DNA fragments are each limited by size. Plasmid sized fragments of up to 10 kb can be easily assembled, but specifically ordered fragments much larger are difficult to assemble by current techniques and require careful handling to avoid breakage. Other types of cloning vehicles allow larger fragments to be cloned and manipulated, but even these have their limits (lambda ~15-20 kb; cosmids ~ 35-40 kb, BAC ~100 kb and YACs ~1000 kb). Therefore, there is a need in the art for the assembly and manipulation of very large DNA fragments.

[6] PCR fragments can be assembled into larger arrangements for useful purposes. This is usually done with the creation of restriction sites in the primer sequences. When the

amplified DNA is cut with a particular restriction enzyme, a short overhang is generated that can be used to assemble two PCR fragments with complementary overhang sequences. However, PCR fragments often have like ends, so that orientation of the resulting fragment is not defined. Further, many restriction enzymes also cleave within a large PCR fragment and cannot be used in this way. When different restriction enzyme sites are used in each of the two amplification primers, the likelihood that one of restriction enzymes will cut within the PCR fragment is doubled. A method that did not depend on restriction enzymes would be of general application and advantageous.

[7] The invention provides a means of assembling PCR fragments that does not rely on restriction enzymes.

SUMMARY OF THE INVENTION

[8] Generally speaking, the invention uses a modified nucleotide at a specific position in the primer which is removed with DNA excision repair enzyme and AP endonuclease or AP lyase. The AP endonuclease or lyase activity may be part of the excision repair enzyme. In one embodiment the modified nucleotide is deoxyuridine and the commercially available enzymes uracil-DNA-glycosylase (4) and T_4 endonuclease V are used to remove the base and cleave the primer to generate a ligatable 5' phosphorylated end.

[9] A wide variety of excision repair/AP endonuclease (or lyase) combinations are known in the art, and some of the excision repair enzymes also exhibit lyase activity, cleaving the DNA backbone on the 3' side of the AP site. Suitable excision repair enzymes include Methyl Purine DNA Glycosylase (recognizes methylated bases), 8-Oxo-Guanine Glycosylase 1 (recognizes 8-oxoG:C pairs and has lyase activity), Endonuclease Three Homolog 1 (recognizes T-glycol, C-glycol, and formamidopyrimidine and has lyase activity), inosine, hypoxanthine-DNA glycosylase; 5-Methylcytosine, 5-Methylcytosine DNA glycosylase; Formamidopyrimidine-DNA-glycosylase (excision of oxidized residue from DNA: hydrolysis of the N-glycosidic bond (DNA glycosylase), beta-elimination (AP-lyase reaction)).

[10] Sequential ligation is facilitated if the first fragment attached to modified streptavidin-coated magnetic beads, or otherwise immobilized. Suitable immobilization methods

include the use of 1) amine-oligos covalently linked to an activated carboxylate group or succinimidyl ester, 2) SH-oligos covalently linked via an alkylating reagent such as an iodoacetamide or maleimide, 3) acrydite-oligos covalently linked through a thioether, 4) antibody-antigen based capture, 5) nucleic acid triplex affinity interaction, 6) immobilized metal affinity chromatography of his-tagged DNA, 7) streptavidin-SBP-Tag, and 8) phenylboronic acid-salicylhydroxamic acid (SHA) based systems, to name a few. Several such fragments can be assembled on the solid substrate to give an in-phase, functional gene. Any substrate format may be employed.

[11] Inclusion of a site specific recombinase site in the first and last PCR fragments of the assembled DNA allows the joined fragments to be removed from the solid assembly system and circularized for transformation. One site specific recombinase system is the cre-lox system. Inclusion of lox sites in the first (5'-most) and last (3'-most) primers, or otherwise within the amplified fragments, allows the enzyme cre to remove and circularize all DNA between the lox sites.

[12] The cre-lox system is the most commonly used site specific recombination system, but the art teaches a very large variety of site specific recombinases that are too numerous to name. Other specific recombinases, including the FlpR, xerD, shufflon, SSV1 integrase, the Tn3 family, the IS6 family, and the lambda integrase-excisionase or terminase/cos reactions, may also be usefully adapted for the invention.

[13] The useful features of this invention are that several fragments can be joined in a defined order allowing the PCR generated DNA fragments to be assembled into an ordered arrangement and formed into a replicating plasmid without use of restriction enzymes. The final product is in a form for transformation into cells and the use of immobilized DNA facilitates the steps and allows for scaling with automated devices.

BRIEF DESCRIPTION OF THE DRAWINGS

[14] Fig. 1. Use of Modified Primers in PCR.

[15] Fig. 2. After PCR the End of the Primer Can be Removed.

[16] Fig. 3. The Protruding Ends can be Made at one of Both Ends of the PCR Fragment.

[17] Fig. 4 A Biotin can be Attached to the End of the DNA Fragment, Then Ligations can be Done Sequentially with the DNA Attached to the Bead. Lox sites are shown in the first and last primers. Treatment with CRE allows removal and circularization of the final assembled product. Also, by including replication and selection functions on the DNA between the lox sites, the DNA will form a functional vector capable of transforming cells.

[18] Fig. 5. Using an Enzyme System to Remove the DNA from the bead and Circularize it.

DESCRIPTION OF EMBODIMENTS OF THE INVENTION

[19] PCR or "polymerase chain reaction" is a techniques that allows the "copying" or "amplification" of a segment of DNA. PCR is routinely used in forensic, medical, and research laboratories. DNA fragments from a few hundred bp to 30,000 bp can be made, and amplifications of fragments less than 5 kb are relatively routine.

[20] The most efficient way to obtain larger fragments (>30 kb) is by PCR amplification using specific primers to precisely define the ends and ligating the component fragments in a defined order. A way to specify the orientation in joining of PCR fragments would allow large defined arrays to be constructed without regard to the source of each PCR fragment. However, the full duplex nature of the PCR fragment does not allow the specificity of joining that can be achieved through annealing of complementary single strand ends.

[21] A number of methods have been used for orientating PCR fragments. These generally involve the use of the extra A present at the 3' end of some PCR fragments, or inclusion of specialized sequences that can be manipulated by restriction enzymes, exonucleases or polymerases to generate a sticky end. Other methods of altering the primer so sticky ends can be formed include placement of an apurinic residue, or spacer not recognized by the PCR polymerase as a 3' blocking agent (1) (e.g., hyperthermophilic polymerases cease extension before a dU residue) (2).

[22] The method that we have developed (3) is based on removal of part of the primer, so that a 3' overhang is produced that conveys specificity in joining to other PCR or vector fragments. The primer is made with a modified nucleotide at a specific position. It is subsequently removed by an excision repair enzyme and the chain cleaved by the action of a AP endonuclease (or lyase).

[23] In our initial work we used deoxyuridine as the modified nucleotide and the commercially available enzyme uracil-DNA-glycosylase (4-6) and T_4 endonuclease V to remove the base and cleave the primer to generate a ligatable 5' phosphorylated end.

[24] After cleavage, the 5' end of the primer is dissociated by heat to produce a protruding 3' end on the PCR fragment. The location of the modified nucleotide within the primer and the specific primer sequence allow a variety of lengths and sequences in the 3' overhang to be produced by this treatment. We have shown several such fragments can be assembled to give an in-phase, functional gene (3).

[25] We have extended the method to allow directed sequential ligation by using a reversible attachment to a solid substrate as a scaffold on which to build the growing DNA fragment. One embodiment uses modified streptavidin-coated magnetic beads, which have been used with single stranded DNA (7). The first PCR fragment is coupled to the beads by conjugating a biotin to the first primer or first PCR fragment. Sequential PCR fragments are added to this growing chain using the excision repair generated overhang system described above.

[26] In order to remove the growing PCR chain from the solid scaffold, we employed the cre-lox system to remove the large assembled fragment and circularize it. Cre has been shown to act on linear DNA in solution (8). The released fragment is in a circular form which allows it to be efficient for transformation as it is removed from the scaffold. In this system, the first PCR fragment contains a lox site, as does the last PCR fragment. The lox sites may be included in the primer, or be located inside (3') of the primer sites. Using the Cre enzyme allows removal and recircularization of the assembled DNA chain from the scaffold.

[27] Additional signals can be incorporated into the assembled DNA fragments, such as an origin of replication, and a selective marker. This allows the circularized fragment to be directly transformed into a cell and selected for. In most embodiments, these signals will be contained in the first or last PCR fragment to avoid interrupting the coding sequences. However, the sites may also be placed between genes or in introns.

EXAMPLE 1.

[28] All PCRs were performed using Gene Amp® reagents (PERKIN-ELMER,™ Norwalk, CT, USA) in 50 µL reaction mixtures with 2.5 U *Taq* DNA polymerase and 2.5 mM $MgCl_2$ in a RoboCycler Temperature Cycler (STRATAGENE,™ La Jolla, CA, U.S.A.). PCR primer sequences are shown in Table 1. Amplifications from pACYC184 (NEW ENGLAND BIOLABS,™ Beverly, MA, U.S.A.) contained 50 ng of plasmid and 40 pmol of each PCR primer CAT3 and CAT5 or SacCAT3 and SacCAT5. Cycling was at 95°C for 1 min. 55°C for 1 min. 72° for 1 min for 30 cycles followed by 1 cycle at 72° C for 3 min. Genomic amplifications from E. *coli* strain W3110 contained 100 ng genomic DNA, 40 pmol of LacST#1 and #2, LacMD#1 and #2, LacEN#1 and #2 (See Table 1.) and were carried out at 95°C for 45s, 55°C for 45s, 72°C for 45s for 30 cycles, then a final extension at 72°C for 3 min.

Table 1. PCR Primer Sequences

CAT3	(5' AGCUCGGCAC GTAAGAGGTT CCAACTTTCA CC 3' [32 nucleotide])
CAT5	(5' AGCUCCAGGC GTTTAAGGGC ACCAATAACT GC 3' [32 nt])
SacCAT3	(5' AGAATGAGCT CCAGGCGTTT AAGGGCACCA ATAACTGC 3' [38 nt])
SacCAT5	(5' TCAATGAGCT CGGCACGTAA GAGGTTCCAA CTTTCACC 3' [38 nt])
LacST#1	(5' AGCUCGCACG ACAGGTTTCC CGACTGGAAA GCGGGC 3' [36 nt])
LacST#2	(5' ACCACCACGC UCATCGATAA TTTCACCGCC G 3' [31 nt])
LacMD#1	(5' AGCGTGGTGG UTATGCCGAT CGCGTCACAC 3' [30 nt])
LacMD#2	(5' AGCGCTGGAU GCGGCGTGCG GTCGGCAAAG 3' [30 nt])
LacEN#1	(5' ATCCAGCGCU GACGGAAGCA AAACACCAGC 3' [30 nt])
LacEN#2	(5' AGCUCAATAC GGGCAGACAT GGCCTGCCCG G 3' [31 nt])

[29] PCR products were generated and purified from pACYC184 using primers CAT3 and CAT5 encoding the entire chloramphenicol acetyltransferase (*cat*) gene. Following PCR, 1 µL (1u) of UDG (LIFE TECHNOLOGIES,™ Gaithersburg, MD, USA) and 8 µL of HAP1 (13 mg/mL) were added, and the PCR products incubated for 30 min at 37°C and 15 min at 65°C,

followed by the addition of spermine (0.2 mM final) with 15-min incubations at 37°C and 15 min at 65°C and 70°C. The enzymatically treated PCR samples were purified with QIAquick® PCR Purification Kit (QIAGEN,™ Chatsworth, CA, USA). Ligations were performed with 200 ng of the treated PCR product and 100 ng of *Sac*I-cleaved dephosphorylated pUC19 (STRATAGENE™) overnight at 16°C. Library-efficiency competent E. *coli* DH5α cells (LITE TECHNOLOGIES™) were transformed according to the manufacturer's protocol and plated on LB plates containing ampicillin (Ap) and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) or ampicillin and chloramphenicol (Cm) and incubated overnight at 37°C. Plasmid DNAs from at least 50 individual colonies from each type of plate were isolated using alkaline lysis, digested with *Sac*I and characterized by electrophoresis using 0.7% agarose gels. In addition, colonies from the Ap/X-gal plates were transferred to Ap/Cm plates.

[30] Alternatively, following UDG treatment, the PCR samples were purified with the QIAquick and then treated with HAP1 in HAP1 buffer (20mM Tris-HCl, pH 8.0, 10 mM NaC1, 1 mM EDTA, 5 mM $MgCl_2$ and 10μg heat-inactivated bovine serum albumin [BSA]) and incubated overnight at room temperature to allow for the spontaneous ß elimination of the 5' dRp group.

[31] Immediately following amplification, the CAT PCR product was treated with 1 μL of UDG, 0.5 μL (10 U) of T4 endonuclease V (EPICENTRE TECHNOLOGIES,™ Madison, WI, USA) and incubated at 37°C for 30 min, then at 70°C for 15 min. The PCR mixture was cleaned with QIAquick. In 20 μL, ligations using 1:1, 1:5 and 1:10 molar ratios of *Sac*I-cleaved, dephosphorylated pUC19 (100 ng) to treated PCR product were incubated overnight at 16°C. Library-efficiency competent E. *coli* DH5α cells were transformed following the manufacturer's protocol and plated on Ap/X-gal or with Ap/Cm plates and incubated overnight at 37°C. Plasmid products from at least 50 individual colonies were isolated and checked as described earlier.

[32] Primers SacCAT3 and SacCAT5 were designed, adding 5 nucleotides to the end from the *Sac*I recognition sequence. The *cat* gene PCR products were purified as before and digested with *Sac*I for 16h at 37°C. Following digestion and enzyme inactivation, the digested PCR fragments were ligated using 1:1, 1:5 and 1:10 molar ratios with the same preparation of pUC19 used above. Library-efficiency competent E. *coli* DH5α cells were transformed and

plated on Ap/X-gal or Ap/Cm plates. Plasmid DNAs from at least 50 colonies were isolated and checked as described.

[33] Using genomic DNA from E *coli* W3110 as template, three PCR products were generated using primers LacST#1 and #2, LacMD#1 and #2 and LacEN#1 and #2. Each PCR product was incubated with UDG and T4 endonuclease V as described. The treated PCR samples were ligated into dephosphorylated *Sac*I-cleaved pSP72 (100 ng) (PROMEGA,™ Madison, WI, USA) at a 1:1 and 1:5 vector/insert molar ratio and transformed into library-efficiency competent E. *coli* DH5α cells, and the cells were plated on LB plates containing ampicillin and X-gal to test for the presence of functional ß-galactosidase. Plasmid DNAs from 50 white and blue colonies were isolated and digested with either *Sac*I, *Hind*III, or *Cla*I and separated by electrophoresis on 0.7% agarose gels. The *Hind*III digest shows total plasmid size, and the C*la*I digest determines insert orientation and integrity of the *lac* operon.

[34] Results are shown in Table 2. and the reader is referred to (3) for additional details.

Table 2.

Product Treatment	**CAT PCR Product* Transformants per µg of Vector**	**LacZ PCR Product* Transformants per µg of Vector**
pUC19 transformation control	3×10^7	
pSP72 control		3×10^7
No treatment	0	0
UDG only, no HAP1	0	0
UDG + HAP1 with no Overnight ß elimination	0	0
UDG + HAP1 with no Overnight ß elimination	4×10^3	
UDG + HAP1 + spermine	4.5×10^3	
UDG + T4 endoV, 1:5 ratio	8×10^5	
Sac digestion, 1:5 ratio Primers with SacI site	1×10^4	
UDG + T4 endoV, 1:1 ratio 3 LacZ fragments		2×10^5
UDG + T4 endoV, 1:1 ratio Entire lac operon		2×10^5
*Data are compiled from at least 3 ligations and transformations.		

EXAMPLE 2.

[35] Our first example demonstrated the viability of the method as used in solution. However, the method was cumbersome, tedious, and not applicable for large scale up. Thus, we now demonstrate a solid phase procedure that is suitable for scale-up and commercial use.

[36] This example was based on the use of a plasmid which, intact, confers resistance to the antibiotic ampicillin and has the lacZ gene which allows for metabolism of the substrate Xgal; an E. coli host with a lacZ-containing plasmid grows as a blue colony on Xgal plates instead of a white colony which the E. coli host produces.

[37] The first step is to design oligonucleotides so that the PCR product A will have an overhang on the 3 end opposite of the biotin and will leave a sticky 3' end after enzymatic treatment with UDG and T4 endonuclease V AND so that there is a biotin attached to the 5 end AND so that the PCR amplified product includes a loxP site near the end where biotin is attached. In our example, this product A also includes the gene which confers resistance to ampicillin.

[38] A second set of primers are designed so that the PCR amplification product B will have a overhang after enzymatic treatment with UNG and T4 endonuclease V which is complementary to the overhang formed on PCR product A, AND so that the PCR product includes a loxP site in an appropriate arrangement near the 3 end.

[39] PCR amplifications were performed to get products A and B and both were treated with T4 DNA polymerase in the presence of ultrapure dNTPs to be certain that the PCR products have been completed and have flush ends. A Wizard cleanup kit is used to remove extra nucleotides and enzyme. Both products are then treated with UNG and T4 endonuclease V to create overhangs. The treated products are now designated as A/U/T for the treated PCR A product and B/U/T for the treated PCR product B.

[40] The treated fragment bearing an attached biotin, A/U/T is bound to DYNALs magnetic dynabeads with attached streptavidin ; the biotin on A/U/T will bind strongly to the streptavidin. As the experiment progresses, the beads magnetic properties are used to quickly wash away reagents and enzymes, change buffer conditions, and add in new reaction

components. This presents a significant commercial advantage over performing the reactions in solution and having to centrifuge the products after every washing step.

[41] After binding a wash is used to remove unbound A/U/T. Then B/U/T is added to the beads which have bound A/U/T. A joining reaction is done to ligate A/U/T and B/U/T so that product, C, is created, and it is bound to the beads. The ligase enzyme is heat inactivated and unreacted B/U/T and ligase reaction components are washed to remove them from the bead mixture. Cre recombinase is added to the beads which have C bound, in order to circularize C using the loxP sites; circularizing detaches C from the beads and forms a product that can be used directly in transformations. Cre recombinase can then be heat inactivated and at this point the DNA can be transformed.

[42] In this example, a restriction enzyme(s) was added which will cut any template DNA (remaining from the PCR reactions) which is contaminating the product C; this step can be used to reduce the number of blue colonies in the transformation step, making identification of white colonies (containing product C) easier. The restriction enzyme(s) are then heat inactivated.

[43] The supernatant from the bead mixture, which contains circularized C, is used to transform E. coli cells. Selection of transformed cells is accomplished by plating the transformed cells on Xgal /Amp plates. White colonies contain the desired fragment C. Blue colonies contain the plasmid used as template for the PCR reactions.

[44] In an experiment performed in 2000, the efficiency of the entire experiment was approximately 100,000 white colonies/ug of A/U/T on the beads. The background level of blue colonies was < 1%. The optimal transformation efficiency for this host cell is 1 billion colonies/ug for simply transforming the cells with a small characterized plasmid.

[45] The reagents used in the bead assembly method:

Reagent	**Source**
Failsafe Polymerase for PCR	EPICENTRE™
T4 DNA polymerase	EPICENTRE™
Ultrapure dNTPs	USB™
PCR preps cleanup kit	WIZARD FROM PROMEGA™
UNG = Uracil N glycosylase	EPICENTRE™
T4 endonuclease V	EPICENTRE™

Reagent	Source
Beads with kilobase binder kit	DYNAL™
Ligase - Fast Link kit	EPICENTRE™
Cre recombinase	INVITROGEN™

[46] All citations are listed here for convenience and each is expressly incorporated by reference in its entirety for any purpose:

1. Gal J, Schnell R, Kalman M. Polymerase dependence of autosticky polymerase chain reaction. Anal Biochem. 2000 Jun 15;282(1):156-8. Gal J, Schnell R, Szekeres S, Kalman M. Directional cloning of native PCR products with preformed sticky ends (autosticky PCR). Mol Gen Genet. 1999 Jan;260(6):569-73.5.
2. Greagg MA, Fogg MJ, Panayotou G, Evans SJ, Connolly BA, Pearl LH. A read-ahead function in archaeal DNA polymerases detects promutagenic template-strand uracil. Proc Natl Acad Sci U S A. 1999 Aug 3;96(16):9045-50.
3. Watson DE, Bennett GN. Cloning and assembly of PCR products using modified primers and DNA repair enzymes. Biotechniques. 1997 Nov;23(5):858-62, 864 (attached as Example 1).
4. Booth PM, Buchman GW, Rashtchian A. Assembly and cloning of coding sequences for neurotrophic factors directly from genomic DNA using polymerase chain reaction and uracil DNA glycosylase. Gene. 1994 Sep 2;146(2):303-8.
5. Rashtchian A, Buchman GW, Schuster DM, Berninger MS. Uracil DNA glycosylase-mediated cloning of polymerase chain reaction-amplified DNA: application to genomic and cDNA cloning. Anal Biochem. 1992 Oct;206(1):91-7.
6. Nisson PE, Rashtchian A, Watkins PC. Rapid and efficient cloning of Alu-PCR products using uracil DNA glycosylase. PCR Methods Appl. 1991 Nov;1(2):120-3.
7. Stahl S, Hansson M, Ahlborg N, Nguyen TN, Liljeqvist S, Lundeberg J, Uhlen M. Solid-phase gene assembly of constructs derived from the Plasmodium falciparum malaria blood-stage antigen Ag332. Biotechniques 1993 Mar;14(3):424-34.
8. Abremski K, Hoess R, Sternberg N. Studies on the properties of P1 site-specific recombination: evidence for topologically unlinked products following recombination. Cell. 1983 Apr;32(4):1301-11

[47] What is claimed is: